數學傳播 49卷4期, pp. 114-124

利用 gradient descent 方法來求近似迴歸直線與近似迴歸方程式

李灃全

前言

利用 gradient descent 方法來求近似迴歸直線與近似迴歸方程式, 並舉實例說明且預測房價, 請看下文的剖析。

本文

一、已知 (x_i, y_i), $i = 1, 2, \ldots, n$, 求迴歸直線

1. 設直線 $y = mx + b$

$$f(m,b) = \frac{1}{2}\sum_{i=1}^{n}[y_i - (mx_i + b)]^2.$$

為了求迴歸直線我們極小化 $f(m,b) = \frac{1}{2}\sum_{i=1}^{n}[y_i - (mx_i + b)]^2$。

2. $\dfrac{\partial f}{\partial m} = \sum_{i=1}^{n}[y_i - (mx_i + b)] \cdot (-x_i) = 0,$

$$\Rightarrow\ -\sum_{i=1}^{n} x_i y_i + m\Big(\sum_{n=1}^{n} x_i^2\Big) + b\Big(\sum_{n=1}^{n} x_i\Big) = 0,$$

$$\Rightarrow\ m\Big(\sum_{i=1}^{n} x_i^2\Big) + bn\bar{x} = \sum_{n=1}^{n} x_i y_i. \tag{1}$$

3. $\dfrac{\partial f}{\partial b} = \sum_{i=1}^{n}[y_i - (mx_i + b)] \cdot (-1) = 0,$

$$\Rightarrow\ \sum_{i=1}^{n} y_i = m\Big(\sum_{n=1}^{n} x_i\Big) + nb\ \Rightarrow\ n\bar{y} = m(n\bar{x}) + nb,$$

$$\Rightarrow\ \bar{y} = m\bar{x} + b. \tag{2}$$

4.

$$\begin{cases} m\Big(\sum_{i=1}^{n} x_i^2\Big) + bn\bar{x} = \sum_{n=1}^{n} x_i y_i, & (1) \\ \bar{y} = m\bar{x} + b. & (2) \end{cases}$$

由 (2) 式可得 $\Rightarrow\ b = \bar{y} - m\bar{x}.$ (3)

代入 (1) 式可得 $m\Big(\sum_{i=1}^{n} x_i^2\Big) + n\bar{x}(\bar{y} - m\bar{x}) = \sum_{n=1}^{n} x_i y_i,$

$$\Rightarrow\ m\Big[\Big(\frac{\sum_{i=1}^{n} x_i^2}{n}\Big) - (\bar{x})^2\Big] = \frac{\sum_{i=1}^{n} x_i y_i}{n} - \bar{x}\cdot\bar{y},$$

$$\Rightarrow\ m = \frac{\frac{\sum_{i=1}^{n} x_i y_i}{n} - \bar{x}\cdot\bar{y}}{\frac{\sum_{i=1}^{n} x_i^2}{n} - (\bar{x})^2}\ \Rightarrow\ m = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\cdot\bar{y}}{\sum_{i=1}^{n} x_i^2 - n(\bar{x})^2}. \quad (4)$$

5. $\dfrac{\partial^2 f}{\partial m^2} = \sum_{i=1}^{n} x_i^2,\ \dfrac{\partial^2 f}{\partial b^2} = \sum_{i=1}^{n} 1 = n,\ \dfrac{\partial^2 f}{\partial m \partial b} = \dfrac{\partial^2 f}{\partial b \partial m} = \sum_{i=1}^{n} x_i$。

$$\text{海森 (Hessian) 行列式} = \Delta = \begin{vmatrix} \frac{\partial^2 f}{\partial m^2} & \frac{\partial^2 f}{\partial m \partial b} \\ \frac{\partial^2 f}{\partial b \partial m} & \frac{\partial^2 f}{\partial b^2} \end{vmatrix} = \begin{vmatrix} \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & n \end{vmatrix}$$

$$= n\sum_{i=1}^{n} x_i^2 - \Big(\sum_{i=1}^{n} x_i\Big)^2 = \sum_{1\le i<j\le n} (x_i - x_j)^2 > 0 \quad \text{且} \quad \frac{\partial^2 f}{\partial m^2} = \sum_{i=1}^{n} x_i^2 > 0.$$

因此當 $m = \dfrac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\cdot\bar{y}}{\sum_{i=1}^{n} x_i^2 - n(\bar{x})^2}$ 且 $b = \bar{y} - m\bar{x}$ 時 $f(m,b) = \dfrac{1}{2}\cdot\sum_{i=1}^{n}[y_i - (mx_i + b)]^2$ 有極小值。

6. 由1. 到 5. 得知, 當 (x_i, y_i), $i = 1, 2, \ldots, n$, 所求迴歸直線爲 $y = mx + b$ 其中

$$m = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\cdot\bar{y}}{\sum_{i=1}^{n} x_i^2 - n(\bar{x})^2}, \quad b = \bar{y} - m\bar{x}.$$

二、例子

已知 $(x_1, y_1) = (1, 5)$, $(x_2, y_2) = (2, 6)$, $(x_3, y_3) = (3, 10)$, $(x_4, y_4) = (4, 13)$, $(x_5, y_5) = (5, 11)$, 求迴歸直線

1. 由一、得知

當 (x_i, y_i), $i = 1, 2, \ldots, 5$, 所求迴歸直線為 $y = mx + b$ 其中

$$m = \frac{\sum\limits_{i=1}^{5} x_i y_i - 5\bar{x} \cdot \bar{y}}{\sum\limits_{i=1}^{5} x_i^2 - 5(\bar{x})^2}, \quad b = \bar{y} - m\bar{x}.$$

2.
$$\sum_{i=1}^{5} x_i = 1 + 2 + 3 + 4 + 5 = 15 \ \Rightarrow\ \bar{x} = \frac{15}{5} = 3,$$
$$\sum_{i=1}^{5} y_i = 5 + 6 + 10 + 13 + 11 = 45 \ \Rightarrow\ \bar{y} = \frac{45}{5} = 9,$$
$$\sum_{i=1}^{5} x_i^2 = 1 + 4 + 9 + 16 + 25 = 55,$$
$$\sum_{i=1}^{5} x_i y_i = 5 + 12 + 30 + 52 + 55 = 154.$$

代入 (4) 式 $m = \dfrac{\sum\limits_{i=1}^{n} x_i y_i - n\bar{x} \cdot \bar{y}}{\sum\limits_{i=1}^{n} x_i^2 - n(\bar{x})^2}$, $n = 5$。得到 $m = \dfrac{\sum\limits_{i=1}^{5} x_i y_i - 5\bar{x} \cdot \bar{y}}{\sum\limits_{i=1}^{5} x_i^2 - 5(\bar{x})^2}$

$= \dfrac{154 - 5(3)(9)}{55 - 5(3)^2} = \dfrac{154 - 135}{55 - 45} = \dfrac{19}{10} = 1.9$。將 $m = 1.9$ 代入 (3) 式

$b = \bar{y} - m\bar{x} = 9 - 1.9(3) = 9 - 5.7 = 3.3$, 因此求得迴歸直線為 $y = 1.9x + 3.3$。

三、利用 gradient descent 方法來求, 二、例子

當 $(x_1, y_1) = (1, 5)$, $(x_2, y_2) = (2, 6)$, $(x_3, y_3) = (3, 10)$, $(x_4, y_4) = (4, 13)$, $(x_5, y_5) = (5, 11)$ 的近似迴歸直線。

1. 起始點 $(m_0, b_0) = (0, 0)$。
2. 機器學習率 $\eta = 0.001$。
3. $(m_{i+1}, b_{i+1}) = (m_i, b_i) - \eta \nabla f(m, b)\Big|_{(m_i, b_i)}$

$$= (m_i, b_i) - \eta\Big(\frac{\partial f(m,b)}{\partial m}\Big|_{(m_i,b_i)}, \frac{\partial f(m,b)}{\partial b}\Big|_{(m_i,b_i)}\Big),$$

$$\text{其中}\quad \frac{\partial f(m,b)}{\partial m}\Big|_{(m_i,b_i)} = \sum_{i=1}^{5}[y_i-(m_ix_i+b_i)]\cdot(-x_i),$$
$$\frac{\partial f(m,b)}{\partial b}\Big|_{(m_i,b_i)} = \sum_{i=1}^{5}[y_i-(m_ix_i+b_i)]\cdot(-1).$$

4. 疊代 10000 次即 $i=0,1,2,\ldots,9999$

得到 $(m_1,b_1)=(0.154,0.045)$, $(m_2,b_2)=(0.298855,0.087465)$,

$(m_3,b_3)=(0.435106,0.127545)$, $(m_4,b_4)=(0.563262,0.165381)$,

$(m_5,b_5)=(0.683802,0.201105)$, $(m_6,b_6)=(0.797176,0.234842)$,

$(m_7,b_7)=(0.903809,0.266710)$, $(m_8,b_8)=(1.004099,0.296820)$,

$(m_9,b_9)=(1.098421,0.325274)$, $(m_{10},b_{10})=(1.187129,0.352171)$,

$$\vdots$$
$$(m_{50},b_{50})=(2.460078,0.796231),$$
$$\vdots$$
$$(m_{100},b_{100})=(2.549805,0.931158),$$
$$\vdots$$
$$(m_{1000},b_{1000})=(2.206322,2.194082),$$
$$\vdots$$
$$(m_{5000},b_{5000})=(1.910405,3.262436),$$
$$\vdots$$
$$(m_{10000},b_{10000})=(1.900152,3.299452).$$

我們求得 $y=m_{10000}x+b_{10000} \;\Rightarrow\; y=1.900152x+3.299452$ 爲近似迴歸直線。

四、多元迴歸分析例子

已知當 $(x_1,t_1,y_1)=(57,8,64)$, $(x_2,t_2,y_2)=(59,10,71)$, $(x_3,t_3,y_3)=(49,6,53)$, $(x_4,t_4,y_4)=(62,11,67)$, $(x_5,t_5,y_5)=(51,8,55)$, $(x_6,t_6,y_6)=(50,7,58)$, $(x_7,t_7,y_7)=(55,10,77)$, $(x_8,t_8,y_8)=(48,9,57)$, 求迴歸方程式 $y=mx+ut+b$.

1. 設迴歸方程式爲 $y=mx+ut+b$, 令

$$f(m,u,b)=\frac{1}{2}\sum_{i=1}^{8}[y_i-(mx_i+ut_i+b)]^2,$$

爲了求迴歸方程式我們極小化

$$f(m,u,b)=\frac{1}{2}\sum_{i=1}^{8}[y_i-(mx_i+ut_i+b)]^2.$$

2. $\dfrac{\partial f}{\partial m}=\displaystyle\sum_{i=1}^{8}[y_i-(mx_i+ut_i+b)]\cdot(-x_i)=0,$

$$\Rightarrow\ -\sum_{i=1}^{8}x_iy_i+m\Big(\sum_{i=1}^{8}x_i^2\Big)+u\Big(\sum_{i=1}^{8}x_it_i\Big)+b\Big(\sum_{i=1}^{8}x_i\Big)=0,$$

$$\Rightarrow m\Big(\sum_{i=1}^{8}x_i^2\Big)+u\Big(\sum_{i=1}^{8}x_it_i\Big)+b\Big(\sum_{i=1}^{8}x_i\Big)=\sum_{i=1}^{8}x_iy_i. \tag{5}$$

3. $\dfrac{\partial f}{\partial u}=\displaystyle\sum_{i=1}^{8}[y_i-(mx_i+ut_i+b)]\cdot(-t_i)=0,$

$$\Rightarrow\ -\sum_{i=1}^{8}t_iy_i+m\Big(\sum_{i=1}^{8}x_it_i\Big)+u\Big(\sum_{i=1}^{8}t_i^2\Big)+b\Big(\sum_{i=1}^{8}t_i\Big)=0,$$

$$\Rightarrow m\Big(\sum_{i=1}^{8}x_it_i\Big)+u\Big(\sum_{i=1}^{8}t_i^2\Big)+b\Big(\sum_{i=1}^{8}t_i\Big)=\sum_{i=1}^{8}t_iy_i. \tag{6}$$

4. $\dfrac{\partial f}{\partial b}=\displaystyle\sum_{i=1}^{8}[y_i-(mx_i+ut_i+b)]\cdot(-1)=0,$

$$\Rightarrow m\Big(\sum_{i=1}^{8}x_i\Big)+u\Big(\sum_{i=1}^{8}t_i\Big)+8b=\sum_{i=1}^{8}y_i. \tag{7}$$

5. 由 (5) (6) (7) 式可得矩陣形式

$$\begin{bmatrix} 8 & \sum_{i=1}^{8}x_i & \sum_{i=1}^{8}t_i \\ \sum_{i=1}^{8}x_i & \sum_{i=1}^{8}x_i^2 & \sum_{i=1}^{8}x_it_i \\ \sum_{i=1}^{8}t_i & \sum_{i=1}^{8}x_it_i & \sum_{i=1}^{8}t_i^2 \end{bmatrix}\begin{bmatrix} b \\ m \\ u \end{bmatrix}=\begin{bmatrix} \sum_{i=1}^{8}y_i \\ \sum_{i=1}^{8}x_iy_i \\ \sum_{i=1}^{8}t_iy_i \end{bmatrix},$$

推得

$$\begin{bmatrix} 8 & 431 & 69 \\ 431 & 23405 & 3762 \\ 69 & 3762 & 615 \end{bmatrix}\begin{bmatrix} b \\ m \\ u \end{bmatrix}=\begin{bmatrix} 502 \\ 27264 \\ 4406 \end{bmatrix},$$

$$\Rightarrow \begin{bmatrix} b \\ m \\ u \end{bmatrix} = \begin{bmatrix} 8 & 431 & 69 \\ 431 & 23405 & 3762 \\ 69 & 3762 & 615 \end{bmatrix}^{-1} \begin{bmatrix} 502 \\ 27264 \\ 4406 \end{bmatrix},$$

$$\Rightarrow \begin{bmatrix} b \\ m \\ u \end{bmatrix} = \frac{1}{13464} \cdot \begin{bmatrix} 241431 & -5487 & 6477 \\ -5487 & 159 & -357 \\ 6477 & -357 & 1479 \end{bmatrix} \cdot \begin{bmatrix} 502 \\ 27264 \\ 4406 \end{bmatrix} = \begin{bmatrix} \frac{1923}{187} \\ \frac{105}{187} \\ \frac{1445}{561} \end{bmatrix},$$

求得迴歸方程式 $y=mx+ut+b=\dfrac{105}{187}x+\dfrac{1445}{561}t+\dfrac{1923}{187}$。

利用 gradient descent 方法來求：

已知當 $(x_1,t_1,y_1)=(57,8,64)$, $(x_2,t_2,y_2)=(59,10,71)$, $(x_3,t_3,y_3)=(49,6,53)$, $(x_4,t_4,y_4)=(62,11,67)$, $(x_5,t_5,y_5)=(51,8,55)$, $(x_6,t_6,y_6)=(50,7,58)$, $(x_7,t_7,y_7)=(55,10,77)$, $(x_8,t_8,y_8)=(48,9,57)$ 時, 近似迴歸方程式爲何?

設近似迴歸方程式爲 $y=mx+ut+b$, $f(m,u,b)=\dfrac{1}{2}\sum_{i=1}^{8}[y_i-(mx_i+ut_i+b)]^2$。

6. 起始點 $(m_0,u_0,b_0)=(0,0,0)$。
7. 機器學習率 $\eta=0.000001$。
8. $(m_{i+1},u_{i+1},b_{i+1})=(m_i,u_i,b_i)-\eta\nabla f(m,u,b)\Big|_{(m_i,u_i,b_i)}$

$$=(m_i,u_i,b_i)-\eta\Big(\frac{\partial f(m,u,b)}{\partial m}\Big|_{(m_i,u_i,b_i)},\frac{\partial f(m,u,b)}{\partial u}\Big|_{(m_i,u_i,b_i)},\frac{\partial f(m,u,b)}{\partial b}\Big|_{(m_i,u_i,b_i)}\Big),$$

其中 $$\frac{\partial f(m,u,b)}{\partial m}\Big|_{(m_i,u_i,b_i)}=\sum_{i=1}^{8}[y_i-(mx_i+ut_i+b)]\cdot(-x_i),$$

$$\frac{\partial f(m,u,b)}{\partial u}\Big|_{(m_i,u_i,b_i)}=\sum_{i=1}^{8}[y_i-(mx_i+ut_i+b)]\cdot(-t_i),$$

$$\frac{\partial f(m,u,b)}{\partial b}\Big|_{(m_i,u_i,b_i)}=\sum_{i=1}^{8}[y_i-(mx_i+ut_i+b)]\cdot(-1).$$

9. 疊代 200000000 次即 $i=0,1,2,\ldots,199999999$, 得到 $(m_{200000000},u_{200000000},b_{200000000})=(0.56150,2.57575,10.28327)$, 因此近似迴歸方程式爲 $y=m_{200000000}x+u_{200000000}t+b_{200000000}$, 即 $y=0.56150x+2.57575t+10.28327$。

五、房屋價格預測

1. x_i 爲第 i 間房子的面積, 單位平方英呎。

   t_i 爲第 i 間房子的臥房數。

   v_i 爲第 i 間房子的屋齡。

   y_i 爲第 i 間房子的價格, 單位美元。

2. 資料

$$\begin{aligned}(x_1,t_1,v_1,y_1) =&(2600,3,20,550000), \quad (x_2,t_2,v_2,y_2) =(3000,4,15,565000),\\ (x_3,t_3,v_3,y_3) =&(3200,3,18,610000), \quad (x_4,t_4,v_4,y_4) =(3600,3,30,595000),\\ (x_5,t_5,v_5,y_5) =&(4000,5,8,760000).\end{aligned}$$

3. 設迴歸方程式爲 $y=mx+ut+rv+b$, 令

$$f(m,u,r,b)=\frac{1}{2}\sum_{i=1}^{5}[y_i-(mx_i+ut_i+rv_i+b)]^2,$$

   爲了求迴歸方程式我們極小化

$$f(m,u,r,b)=\frac{1}{2}\sum_{i=1}^{5}[y_i-(mx_i+ut_i+rv_i+b)]^2.$$

4. $\dfrac{\partial f}{\partial m}=\sum_{i=1}^{5}[y_i-(mx_i+ut_i+rv_i+b)]\cdot(-x_i)=0,$

$$\Rightarrow\ -\sum_{i=1}^{5}x_iy_i+m\Big(\sum_{i=1}^{5}x_i^2\Big)+u\Big(\sum_{i=1}^{5}x_it_i\Big)+r\Big(\sum_{i=1}^{5}x_iv_i\Big)+b\Big(\sum_{i=1}^{5}x_i\Big)=0,$$

$$\Rightarrow m\Big(\sum_{i=1}^{5}x_i^2\Big)+u\Big(\sum_{i=1}^{5}x_it_i\Big)+r\Big(\sum_{i=1}^{5}x_iv_i\Big)+b\Big(\sum_{i=1}^{5}x_i\Big)=\sum_{i=1}^{5}x_iy_i. \tag{8}$$

5. $\dfrac{\partial f}{\partial u}=\sum_{i=1}^{5}[y_i-(mx_i+ut_i+rv_i+b)]\cdot(-t_i)=0,$

$$\Rightarrow\ -\sum_{i=1}^{5}t_iy_i+m\Big(\sum_{i=1}^{5}x_it_i\Big)+u\Big(\sum_{i=1}^{5}t_i^2\Big)+r\Big(\sum_{i=1}^{5}t_iv_i\Big)+b\Big(\sum_{i=1}^{5}t_i\Big)=0,$$

$$\Rightarrow m\Big(\sum_{i=1}^{5}x_it_i\Big)+u\Big(\sum_{i=1}^{5}t_i^2\Big)+r\Big(\sum_{i=1}^{5}t_iv_i\Big)+b\Big(\sum_{i=1}^{5}t_i\Big)=\sum_{i=1}^{5}t_iy_i. \tag{9}$$

6. $\dfrac{\partial f}{\partial r} = \sum_{i=1}^{5}[y_i - (mx_i + ut_i + rv_i + b)] \cdot (-v_i) = 0,$

$$\Rightarrow -\sum_{i=1}^{5} v_i y_i + m\Big(\sum_{i=1}^{5} v_i x_i\Big) + u\Big(\sum_{i=1}^{5} v_i t_i\Big) + r\Big(\sum_{i=1}^{5} v_i^2\Big) + b\Big(\sum_{i=1}^{5} v_i\Big) = 0,$$

$$\Rightarrow m\Big(\sum_{i=1}^{5} v_i x_i\Big) + u\Big(\sum_{i=1}^{5} v_i t_i\Big) + r\Big(\sum_{i=1}^{5} v_i^2\Big) + b\Big(\sum_{i=1}^{5} v_i\Big) = \sum_{i=1}^{5} v_i y_i. \tag{10}$$

7. $\dfrac{\partial f}{\partial b} = \sum_{i=1}^{5}[y_i - (mx_i + ut_i + rv_i + b)] \cdot (-1) = 0,$

$$\Rightarrow -\sum_{i=1}^{5} y_i + m\Big(\sum_{i=1}^{5} x_i\Big) + u\Big(\sum_{i=1}^{5} t_i\Big) + r\Big(\sum_{i=1}^{5} v_i\Big) + 5b = 0,$$

$$\Rightarrow m\Big(\sum_{i=1}^{5} x_i\Big) + u\Big(\sum_{i=1}^{5} t_i\Big) + r\Big(\sum_{i=1}^{5} v_i\Big) + 5b = \sum_{i=1}^{5} y_i. \tag{11}$$

8. 由 (8) (9) (10) (11) 式可得矩陣形式

$$\begin{bmatrix} 5 & \sum_{i=1}^{5} x_i & \sum_{i=1}^{5} t_i & \sum_{i=1}^{5} v_i \\ \sum_{i=1}^{5} x_i & \sum_{i=1}^{5} x_i^2 & \sum_{i=1}^{5} t_i x_i & \sum_{i=1}^{5} v_i x_i \\ \sum_{i=1}^{5} t_i & \sum_{i=1}^{5} t_i x_i & \sum_{i=1}^{5} t_i^2 & \sum_{i=1}^{5} v_i t_i \\ \sum_{i=1}^{5} v_i & \sum_{i=1}^{5} v_i x_i & \sum_{i=1}^{5} v_i t_i & \sum_{i=1}^{5} v_i^2 \end{bmatrix} \begin{bmatrix} b \\ m \\ u \\ r \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{5} y_i \\ \sum_{i=1}^{5} x_i y_i \\ \sum_{i=1}^{5} t_i y_i \\ \sum_{i=1}^{5} v_i y_i \end{bmatrix},$$

推得

$$\begin{bmatrix} 5 & 16400 & 18 & 91 \\ 16400 & 54960000 & 60200 & 294600 \\ 18 & 60200 & 68 & 304 \\ 91 & 294600 & 304 & 1913 \end{bmatrix} \begin{bmatrix} b \\ m \\ u \\ r \end{bmatrix} = \begin{bmatrix} 3080000 \\ 10259000000 \\ 11325000 \\ 54385000 \end{bmatrix},$$

$$\Rightarrow \begin{bmatrix} b \\ m \\ u \\ r \end{bmatrix} = \begin{bmatrix} 5 & 16400 & 18 & 91 \\ 16400 & 54960000 & 60200 & 294600 \\ 18 & 60200 & 68 & 304 \\ 91 & 294600 & 304 & 1913 \end{bmatrix}^{-1} \begin{bmatrix} 3080000 \\ 10259000000 \\ 11325000 \\ 54385000 \end{bmatrix},$$

$$\Rightarrow \begin{bmatrix} b \\ m \\ u \\ r \end{bmatrix} = \frac{1}{640000000} \cdot \begin{bmatrix} 18613240000 & 732400 & -3843160000 & -387480000 \\ 732400 & 1324 & -1031600 & -74800 \\ -3843160000 & -1031600 & 1424440000 & 115320000 \\ -387480000 & -74800 & 115320000 & 11960000 \end{bmatrix}$$

$$\cdot \begin{bmatrix} 3080000 \\ 10259000000 \\ 11325000 \\ 54385000 \end{bmatrix} = \begin{bmatrix} 383725 \\ 137.25 \\ -26025 \\ -6825 \end{bmatrix},$$

求得迴歸方程式 $y=mx+ut+rv+b = 137.25x - 26025t - 6825v + 383725$。

(1) 將 $(x,t,v) = (3000,3,40)$ 代入迴歸方程式

$$\begin{aligned} y &= 137.25x - 26025t - 6825v + 383725 \\ &= 137.25(3000) - 26025(3) - 6825(40) + 383725 = 444400. \end{aligned}$$

我們可以預測當房子的面積 3000 平方英呎, 有 3 間臥房, 屋齡 40 年, 房子的價格爲 444400 美元。

(2) 將 $(x,t,v) = (2500,4,5)$ 代入迴歸方程式

$$\begin{aligned} y &= 137.25x - 26025t - 6825v + 383725 \\ &= 137.25(2500) - 26025(4) - 6825(5) + 383725 = 588625. \end{aligned}$$

我們可以預測當房子的面積 2500 平方英呎, 有 4 間臥房, 屋齡 5 年, 房子的價格爲 588625 美元。

利用 gradient descent 方法來求：

已知當 $(x_1,t_1,v_1,y_1) = (2600,3,20,550000)$, $(x_2,t_2,v_2,y_2) = (3000,4,15,565000)$,
$(x_3,t_3,v_3,y_3) = (3200,3,18,610000)$, $(x_4,t_4,v_4,y_4) = (3600,3,30,595000)$,
$(x_5,t_5,v_5,y_5) = (4000,5,8,760000)$

時, 近似迴歸方程式爲何?

設近似迴歸方程式爲 $y=mx+ut+rv+b$, 令

$$f(m,u,r,b) = \frac{1}{2}\sum_{i=1}^{5}[y_i-(mx_i+ut_i+rv_i+b)]^2,$$

1. 起始點 $(m_0,u_0,v_0,b_0) = (0,0,0,0)$。

2. 機器學習率 $\eta = 0.01$。
3. $(m_{i+1}, u_{i+1}, r_{i+1}, b_{i+1}) = (m_i, u_i, r_i, b_i) - \eta\nabla f(m,u,r,b)\Big|_{(m_i,u_i,r_i,b_i)}$

$$= (m_i, u_i, r_i, b_i) - \eta\Big(\frac{\partial f(m,u,r,b)}{\partial m}\Big|_{(m_i,u_i,r_i,b_i)}, \frac{\partial f(m,u,r,b)}{\partial u}\Big|_{(m_i,u_i,r_i,b_i)}, \frac{\partial f(m,u,r,b)}{\partial r}\Big|_{(m_i,u_i,r_i,b_i)}, \frac{\partial f(m,u,r,b)}{\partial b}\Big|_{(m_i,u_i,r_i,b_i)}\Big),$$

其中 $\frac{\partial f(m,u,r,b)}{\partial m}\Big|_{(m_i,u_i,r_ib_i)} = \sum_{i=1}^{5}[y_i - (mx_i + ut_i + rv_i + b)] \cdot (-x_i)$

$$\frac{\partial f(m,u,r,b)}{\partial u}\Big|_{(m_i,u_i,r_i,b_i)} = \sum_{i=1}^{5}[y_i - (mx_i + ut_i + rv_i + b)] \cdot (-t_i)$$

$$\frac{\partial f(m,u,r,b)}{\partial r}\Big|_{(m_i,u_i,r_i,b_i)} = \sum_{i=1}^{5}[y_i - (mx_i + ut_i + rv_i + b)] \cdot (-v_i)$$

$$\frac{\partial f(m,u,r,b)}{\partial b}\Big|_{(m_i,u_i,r_i,b_i)} = \sum_{i=1}^{5}[y_i - (mx_i + ut_i + rv_i + b)] \cdot (-1).$$

4. 將 5 組資料

$(x_1, t_1, v_1, y_1) = (2600, 3, 20, 550000)$, $(x_2, t_2, v_2, y_2) = (3000, 4, 15, 565000)$,
$(x_3, t_3, v_3, y_3) = (3200, 3, 18, 610000)$, $(x_4, t_4, v_4, y_4) = (3600, 3, 30, 595000)$,
$(x_5, t_5, v_5, y_5) = (4000, 5, 8, 760000)$, 標準化。

5. 疊代 90000 次即 $i = 0, 1, 2, \ldots, 89999$, 得到 $m = 0.883772$, $u = -0.277378$, $r = -0.651639$, $b = 0.000000$, 因此標準化後近似迴歸方程式爲

$$y_{std} = 0.883772x_{std} - 0.277378t_{std} - 0.651639v_{std} + 0.000000.$$

最終回歸參數 (原始尺度), 得到近似迴歸方程式爲

$$y = 137.25x - 26025t - 6825v + 383725.$$

結論

1. 當 (x_i, y_i), $i = 1, 2, \ldots, n$ 迴歸直線爲 $y = mx + b$ 其中

$$m = \frac{\sum\limits_{i=1}^{n} x_i y_i - n\bar{x}\cdot\bar{y}}{\sum\limits_{i=1}^{n} x_i^2 - n(\bar{x})^2}, \qquad b = \bar{y} = m\bar{x}.$$

2. $(x_1,y_1)=(1,5)$, $(x_2,y_2)=(2,6)$, $(x_3,y_3)=(3,10)$, $(x_4,y_4)=(4,13)$, $(x_5,y_5)=(5,11)$, 利用公式解, 求得迴歸直線爲 $y=1.9x+3.3$。
3. $(x_1,y_1)=(1,5)$, $(x_2,y_2)=(2,6)$, $(x_3,y_3)=(3,10)$, $(x_4,y_4)=(4,13)$, $(x_5,y_5)=(5,11)$, 利用 gradient descent 方法, 起始點 $(m_0,b_0)=(0,0)$, 機器學習率 $\eta=0.001$, $(m_{i+1},b_{i+1})=(m_i,b_i)-\eta\nabla f(m_i,b_i)$, 疊代 10000 次, 求得近似迴歸直線爲 $y=m_{10000}x+b_{10000}$ 即 $y=1.900152x+3.299452$。
4. $(x_1,y_1)=(1,5)$, $(x_2,y_2)=(2,6)$, $(x_3,y_3)=(3,10)$, $(x_4,y_4)=(4,13)$, $(x_5,y_5)=(5,11)$, 利用 gradient descent 方法, 起始點 $(m_0,b_0)=(0,0)$, 機器學習率 $\eta=0.001$, $(m_{i+1},b_{i+1})=(m_i,b_i)-\eta\nabla f(m_i,b_i)\Big|_{m_i,b_i}$, 爲了求更接近迴歸直線, 我們疊代 20000 次, 有效位數小數點後 6 位, 求得近似迴歸直線爲 $y=m_{20000}x+b_{20000}=1.900000x+3.300000$ 與公式解求得迴歸直線 $y=1.9x+3.3$ 相等。
5. 已知當 $(x_1,t_1,y_1)=(57,8,64)$, $(x_2,t_2,y_2)=(59,10,71)$, $(x_3,t_3,y_3)=(49,6,53)$, $(x_4,t_4,y_4)=(62,11,67)$, $(x_5,t_5,y_5)=(51,8,55)$, $(x_6,t_6,y_6)=(50,7,58)$, $(x_7,t_7,y_7)=(55,10,77)$, $(x_8,t_8,y_8)=(48,9,57)$, 求得迴歸方程式爲 $y=mx+ut+b=\dfrac{105}{187}x+\dfrac{1445}{561}t+\dfrac{1923}{187}$, 利用 gradient descent 方法, 求得近似迴歸方程式爲 $y=m_{200000000}x+u_{200000000}t+b_{200000000}$, 即 $y=0.56150x+2.57575t+10.28327$ 其中 $m=\dfrac{105}{187}\approx 0.561497326$, $u=\dfrac{1445}{561}\approx 2.575757576$, $b=\dfrac{1923}{187}\approx 10.283422459.$
6. 給 5 組資料 (x_1,t_1,v_1,y_1)=$(2600,3,20,550000)$, (x_2,t_2,v_2,y_2)=$(3000,4,25,565000)$, $(x_3,t_3,v_3,y_3)=(3200,3,28,610000)$, $(x_4,t_4,v_4,y_4)=(3600,3,30,595000)$, $(x_5,t_5,v_5,y_5)=(4000,5,8,760000)$, 求得迴歸方程式 $y=mx+ut+rv+b=137.25x-26025t-6825v+383725$ 並預測 2 組房價。同時利用 gradient descent 方法, 標準化 5 組資料後求得近似迴歸方程式爲

$$y_{std}=0.883772x_{std}-0.277378t_{std}-0.651639v_{std}+0.000000.$$

最終回歸參數 (原始尺度), 得到近似迴歸方程式爲

$$y=137.25x-26025t-6825v+383725.$$

參考文獻

1. 梯度下降法 (Gradient descent), https://zh.wikipedia.org/zh-tw/梯度下降法

—本文作者投稿時爲國立宜蘭大學人工智慧研究所二年級學生—